IFW/DAC #

PATENT
Customer Number 22,852
Attorney Docket No. 08702.0079-03000

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	)	
	)	
CELESTE et al.	)	Group Art Unit: 1642
	)	
Application No.: 10/779,635	)	Examiner: Elizabeth Kemmerer
	)	
Filed: February 18, 2004	)	
	)	
For: TENDON-INDUCING COMPOSITIONS	)	
	)	

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

PETITION TO EXPUNGE INFORMATION UNDER 37 C.F.R. § 1.59(b)

Applicants hereby petition under 37 C.F.R. § 1.59(b) to expunge the Form PTO-1449 initialed by the Examiner on July 26, 2004, and returned with the Office Action of July 27, 2004 (copy enclosed).

Applicants submit that, on information and belief:

(a) this Form PTO-1449 was not submitted to the Patent Office in this application;

(b) this Form PTO-1449 was submitted to the Patent Office in another application filed on the same day this application was filed (February 18, 2004); and

(c) this Form PTO-1449 was erroneously included in this application's file wrapper and should be expunged from the record.

10/27/2004 SDENBOB1 00000023 10779635
01 FC:1460 130.00 OP

This petition is accompanied by a fee of $130.00, as required under 37 C.F.R. §§ 1.17(h) and 1.59(b).

Please grant any extensions of time required to enter this response and charge any additional required fees to our Deposit Account No. 06-0916.

Respectfully submitted,

FINNEGAN, HENDERSON, FARABOW,
GARRETT & DUNNER, L.L.P.

Dated: 10/25/04

By: Elizabeth McNamee
Elizabeth E. McNamee
Reg. No. 54,696

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/779,635	02/18/2004	Anthony J. Celeste	08702-0079-03000	9582

22852 7590 07/27/2004

FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER LLP
1300 I STREET, NW
WASHINGTON, DC 20005

EXAMINER	
KEMMERER, ELIZABETH	
ART UNIT	**PAPER NUMBER**
1646	

DATE MAILED: 07/27/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Docketed 7/28/04 Attorney LDM/MM/EEM
Case 08702-0079-03
Due Date 10-27-04
Action Resp due
By KO

PTO-90C (Rev. 10/03)

Office Action Summary	Application No.	Applicant(s)	
	10/779,635	CELESTE ET AL.	
	Examiner	Art Unit	
	Elizabeth C. Kemmerer, Ph.D.	1646	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *18 February 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *18 February 2004* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *2/18/04*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Status of Application, Amendments, And/Or Claims

The preliminary amendment of 18 February 2004 has been entered in full. Claims 2-28 are canceled. Claim 1 is under examination.

Sequence Rules

The instant application is not fully in compliance with the sequence rules, 37 CFR 1.821-1.825, because each disclosure of a sequence encompassed by the rules is not accompanies by the required reference to the relevant sequence identifier (SEQ ID NO). Specifically, the sequences in the drawings are not accompanied by reference to SEQ ID NOS:. Correction is required.

35 U.S.C. § 112, First Paragraph

The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claim 1 is rejected under 35 U.S.C. 112, first paragraph, because the specification, while being enabling for the claimed DNA wherein the recited BMP-12 related protein is chosen from BMP-12 and MP52, and wherein the protein has the ability to induce the formation of tendon/ligament-like tissue, does not reasonably provide enablement for DNA molecules encoding other BMP-12 like proteins. The specification does not enable any person skilled in the art to which it pertains, or with

which it is most nearly connected, to make and/or use the invention commensurate in scope with these claims.

The claims are directed to DNA molecules encoding a BMP-12 related protein. The specification defines a BMP-12 related protein as follows:

> BMP-12 related proteins are a subset of the BMP/TGF-β/Vg-I family of proteins, including BMP-12 and VL-1, which are defined as tendon/ligament-like tissue inducing proteins encoded by DNA sequences which are cloned and identified, e.g., using PCR, using BMP-12 specific primers, such as primers #6 and #7 described below, with reduced stringency conditions. It is preferred that the DNA sequences encoding BMP-12 related proteins share at least about 80% homology at the amino acid level from amino acids with amino acids #3 to #103 of SEQ ID NO:1. The DNA molecules preferably have a DNA sequence encoding the BMP-12 protein, the sequence of which is provided in SEQ ID NO:1, or a BMP-12 related protein as further described herein. Both the BMP-12 protein and BMP-12 related proteins are characterized by the ability to induce the formation of tendon/ligament-like tissue in the assay described in the examples.

Thus, BMP-12 related proteins are very loosely defined structurally. Reduced stringency conditions allow for cloning of distantly related sequences. The encoded protein must have an activity. However, the instant fact pattern is similar to that in *In re Hyatt*, 708 F.2d 712, 218 USPQ 195 (Fed. Cir. 1983), wherein a single means claim which covered every conceivable means for achieving the stated purpose was held nonenabling for the scope of the claim because the specification at most disclosed only those means known to the inventors. When claims depend on a recited property, a fact situation comparable to *Hyatt* is possible, where the claim covers every conceivable structure (means) for achieving the stated property (result) while the specification discloses at most only those known to the inventor. See also *Fiers v. Sugano*, 984 F.2d 164, 25 USPQ2d 1601 (Fed. Cir. 1993), and MPEP § 2164.08(a). The specification

discloses sequences for BMP-12 and MP52. These proteins are highly homologous. It is clear that there are sequences which are required for the tendon/ligament inducing activity. However, the claims fail to recite any structural limitations, and thus the skilled artisan would have to resort to trial and error experimentation to identify compounds meeting the functional limitations of the claims, even though a suitable assay for tendon/ligament inducing activity is disclosed. At the time of the invention, the state of the art established that mutation of naturally occurring sequences was more likely than not to result in loss of biological activity. While it is known that many amino acid substitutions are generally possible in any given protein the positions within the protein's sequence where such amino acid substitutions can be made with a reasonable expectation of success are limited. Certain positions in the sequence are critical to the protein's structure/function relationship, e.g. such as various sites or regions directly involved in binding, catalysis and in providing the correct three-dimensional spatial orientation of binding and catalytic sites. These or other regions may also be critical determinants of antigenicity. These regions can tolerate only relatively conservative substitutions or no substitutions. See Ngo et al. and Wells.

Due to the large quantity of experimentation necessary to determine what structural features are requires for tendon/ligament inducing activity, the lack of direction/guidance presented in the specification regarding muteins of naturally occurring BMP-12 related proteins, the absence of working examples directed to same, the complex nature of the invention, the state of the prior art which establishes the unpredictability of the effects of mutation on protein structure and function, and the

breadth of the claims which fail to recite any structural limitations, undue experimentation would be required of the skilled artisan to make and/or use the claimed invention in its full scope.

Claim 1 is rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

The claim is directed to a genus of DNA molecules that encode a generically defined BMP-12 related protein.

To provide evidence of possession of a claimed genus, the specification must provide sufficient distinguishing identifying characteristics of the genus. The factors to be considered include disclosure of compete or partial structure, physical and/or chemical properties, functional characteristics, structure/function correlation, methods of making the claimed product, or any combination thereof. In this case, the only factor present in the claim is a partial structure in the form of a recitation of percent identity. There is not even identification of any particular portion of the structure that must be conserved. Accordingly, in the absence of sufficient recitation of distinguishing identifying characteristics, the specification does not provide adequate written description of the claimed genus.

Vas-Cath Inc. v. Mahurkar, 19USPQ2d 1111, clearly states that "applicant must convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession *of the invention*. The invention is, for purposes of the 'written description' inquiry, *whatever is now claimed.*" (See page 1117.) The specification does not "clearly allow persons of ordinary skill in the art to recognize that [he or she] invented what is claimed." (See *Vas-Cath* at page 1116).

With the exception of BMP-12 and MP52, the skilled artisan cannot envision the detailed chemical structure of the encompassed polypeptides, and therefore conception is not achieved until reduction to practice has occurred, regardless of the complexity or simplicity of the method of isolation. Adequate written description requires more than a mere statement that it is part of the invention and reference to a potential method of isolating it. The compound itself is required. See *Fiers v. Revel*, 25 USPQ2d 1601 at 1606 (CAFC 1993) and *Amgen Inc. v. Chugai Pharmaceutical Co. Ltd.*, 18 USPQ2d 1016.

One cannot describe what one has not conceived. See *Fiddes v. Baird*, 30 USPQ2d 1481 at 1483. In *Fiddes*, claims directed to mammalian FGF's were found to be unpatentable due to lack of written description for that broad class. The specification provided only the bovine sequence.

Therefore, only DNA encoding isolated polypeptides comprising the amino acid sequence set forth in BMP-12 and MP52, but not the full breadth of the claim meets the written description provision of 35 U.S.C. §112, first paragraph. Applicant is reminded

that *Vas-Cath* makes clear that the written description provision of 35 U.S.C. §112 is severable from its enablement provision (see page 1115).

Double Patenting

The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. See *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970);and, *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the conflicting application or patent is shown to be commonly owned with this application. See 37 CFR 1.130(b).

Effective January 1, 1994, a registered attorney or agent of record may sign a terminal disclaimer. A terminal disclaimer signed by the assignee must fully comply with 37 CFR 3.73(b).

Claim 1 is rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 1 and 11-15 of U.S. Patent No. 6,027,919. Although the conflicting claims are not identical, they are not patentably distinct from each other because the patented claims are multiple species of the instantly claimed genus. Since the patented species claims in part define the claimed genus, they render obvious the claimed genus.

Claim 1 is rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 1 and 2 of U.S. Patent No. 6,284,872. Although the conflicting claims are not identical, they are not patentably

distinct from each other because the patented claims are directed to a subgenus of the instantly claimed DNA molecules. The patented claims specifically recite the preferred embodiments of the instantly claimed genus, specifically, BMP-12 and MP52. Therefore, the pending claims are rendered obvious by the recited species of the patented claims.

Conclusion

No claims are allowed.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Elizabeth C. Kemmerer, Ph.D. whose telephone number is (571) 272-0874. The examiner can normally be reached on Monday through Thursday, 7:00 a.m. to 5:30 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Brenda Brumback, Ph.D., can be reached on (571) 272-0961. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

ECK

Elizabeth C. Kemmerer

ELIZABETH KEMMERER
PRIMARY EXAMINER

OMB No. 0651-0011

INFORMATION DISCLOSURE CITATION

Atty. Docket No. 04329.3243	Application No.
Applicant Hirofumi KANAI	
Filing Date February 18, 2004	Group:

U.S. PATENT DOCUMENTS

Examiner Initial*		Document Number	Issue Date	Name	Class	Sub Class	Filing Date If Appropriate

FOREIGN PATENT DOCUMENTS

		Document Number	Publication Date	Country	Class	Sub Class	Translation Yes or No
CH		2001-242965	09/07/2001	Japan	—	—	No

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

Examiner E. Kennemer	Date Considered 11/26/04

*Examiner: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

Form PTO 1449 — Patent and Trademark Office - U.S. Department of Commerce

Notice of References Cited	Application/Control No. 10/779,635	Applicant(s)/Patent Under Reexamination CELESTE ET AL.	
	Examiner Elizabeth C. Kemmerer, Ph.D.	Art Unit 1646	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-6,027,919	02-2000	Celeste et al.	435/69.7
*	B	US-6,284,872	09-2001	Celeste et al.	530/399
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Ngo et al., 1994, The Protein Folding Problem and tertiary Structure Prediction, Merz et al., eds., Birkhauser, Boston, pp. 492-495.
	V	Wells, 1990, Biochemistry 29:8509-8517.
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

14

Computational Complexity, Protein Structure Prediction, and the Levinthal Paradox

J. Thomas Ngo, Joe Marks, and Martin Karplus

1. Perspectives and Overview

A protein molecule is a covalent chain of amino acid residues. Although it is topologically linear, in physiological conditions it folds into a unique (though flexible) three-dimensional structure. This structure, which has been determined by x-ray crystallography and nuclear magnetic resonance for many proteins (Bernstein et al., 1977; Abola et al., 1987), is referred to as the native structure. As demonstrated by the experiments of Anfinsen and co-workers (Anfinsen et al., 1961; Anfinsen, 1973), at least some protein molecules, when denatured (unfolded) by disrupting conditions in their environment (such as acidity or high temperature) can spontaneously refold to their native structures when proper physiological conditions are restored. Thus, all of the information necessary to determine the native structure can be contained in the amino acid sequence.

From this observation, it is reasonable to suppose that the native fold of a protein can be predicted *computationally* using information only about its chemical composition. In particular, it should be possible to write down a mathematical problem that, when solved, gives the native conformation of the protein. This procedure would be self-contained, in the sense that no additional information about the biology of protein synthesis would be required. Further, it is reasonable to hope that this procedure could be accomplished without requiring an astronomical amount of computer resources, given the observation that polypeptide chains do fold to their

The Protein Folding Problem and Tertiary Structure Prediction
K. Merz, Jr. and S. Le Grand, Editors

which depends steeply on the energy gap U. Given the assumptions that $N = 100$ and $n = 2$, it was found that in the limit $U \to 0$, the first-passage time is nearly 10^{30} years. However, a modest change to the value of U, say $U = 2kT$, lowers the first-passage time to under one second. (The base of the exponential, $1 + n\exp(-U/kT)$, is equal to 3 when $U = 0$, but 1.27 when $U = 2kT$.)

The analysis of Zwanzig et al. resolves a form of the Levinthal paradox in which the absence of clues about the form of the native state is the sole basis for expecting exponential-time folding. However, it does not resolve the form of the paradox based on computational complexity, since the optimization problem implied by the underlying model can be solved trivially in linear time. The reason for the tractability of the underlying model is the lack of long-range interactions, which are critical to rendering PSP NP-hard (Ngo and Marks, 1992), and essential for cooperativity (Karplus and Shakhnovich, 1992).

6. Future Work

It is not known whether there exists an efficient algorithm for predicting the structure of a given protein from its amino acid sequence alone. Decades of research have failed to produce such an algorithm, yet Nature seems to solve the problem. Proteins do fold! The "direct" approach to structure prediction, that of directly simulating the folding process, is not yet possible because contemporary hardware falls eight to nine orders of magnitude short of the task. However, while this difference is large, it is not astronomical. Would this "direct" approach constitute an efficient and correct algorithm for protein-structure prediction? Too little is known about protein folding, and about the future of computing technology, to be able to answer this question at this time.

The results reviewed here (Section 3) do not completely rule out the existence of a protein-structure prediction algorithm that is both efficient and correct, in the precise senses of those words used throughout this chapter. In particular, it remains formally possible that there is a restricted form of PSP that is efficiently solvable, but subsumes protein-structure prediction. How can this possibility be investigated?

A standard strategy in the analysis of any NP-hard problem is to examine restricted forms of the problem systematically, classifying each as tractable or NP-hard, and thereby exposing the sources of the complexity. Barahona's results with Ising spin-glass models, which were described briefly in Section 4, are exemplary of this approach. While the particular

restrictions chosen by Barahona for spin glasses (reduction of dimensionality and removal of the magnetic field) are not suitable for protein-structure prediction, the overall strategy of examining restricted forms is appropriate. Some restricted form of PSP in which compactness plays a critical role is a candidate for this type of analysis (Section 4.6).

The approach of considering restricted forms has worked well for dozens of important problems that are relatively "clean" and abstract (Garey and Johnson, 1979), but it may be difficult to pursue in the case of protein-structure prediction. In the former case, the problem shown to be NP-hard is usually as general as would actually be required in practice. In the latter case, what is desired is not an algorithm that can handle all possible instances of PSP (Section 3), but merely one that works for proteins. Thus, the fact that PSP is a generalization of protein-structure prediction makes the result that PSP is NP-hard less limiting than it could be.

Ideally, one would like to demonstrate the NP-hardness of a problem that is more *specific*, not more general, than protein-structure prediction, because that would automatically prove the NP-hardness of protein-structure prediction itself. This would entail finding an efficient transformation from some existing NP-complete problem that generates instances of PSP that are proteins by every conceivable criterion.[38] It is difficult to see how such a transformation might proceed.[39]

An alternative approach that may be nearly as instructive is to use the currently available result regarding PSP as a baseline in a continuing comparative analysis—to find restricted forms of PSP that are NP-hard but as specialized as possible, and to find others that are tractable but as general as possible. The motivations for pursuing this methodology are both practical and theoretical:

- Every NP-hardness result permits us to know in advance that a certain group of algorithms is likely to fail, and is therefore not worth pursuing (Section 4).
- Conversely, every NP-hardness result helps identify a source of complexity in protein-structure prediction, and therefore what must be stripped away from the problem before it is reasonable to attempt efficient solution.

  The work of Finkelstein and Reva (1992) is a good example: an approach to structure prediction with a guaranteed polynomial time bound was developed. The critical assumption behind the algorithm is that only nonbonded interactions between nearest neighbors along the chain are significant. Because of this assumption, the algorithm cannot solve all instances of PSP, but instead is restricted to instances in

which only nonbonded interactions between nearest neighbors along the chain are nonzero.[40] This violates the requirements of the reduction from Partition to PSP, in which nonbonded interactions between sites distant from each other along the chain are essential. Thus, the problem is similar in character to that examined by Zwanzig et al. (Section 5.3). While the Finkelstein–Reva algorithm was not inspired by an NP-hardness result, the underlying strategy is similar to how NP-hardness results might be used; they removed from the problem what they observed to be a source of complexity. However, in this case, removing the source of complexity led to a problem different from that posed by protein folding, in which long-range interactions play an essential role.

- The NP-hardness of PSP serves as the premise for a reformulation of the Levinthal paradox (Section 5), whose conventional form is based on a model of folding that is in conflict with known experimental results. A motivation for pursuing an analysis of the computational complexity of protein-structure prediction is to assist in the constructive role of the Levinthal paradox—to help focus attention on the key questions in protein folding.

  A small number of reasonably well-defined potential resolutions to the computational-complexity form of the Levinthal paradox were listed in Section 5. One of the possible resolutions is that protein-structure prediction is tractable. NP-hardness results with restricted forms of PSP would make that possible resolution less likely, thus lending credence to the alternatives.

  Attempts to resolve the Levinthal paradox, which play a valid and useful role in helping to understand how proteins fold, can lead to confusion because the premises of the original form of the paradox are not well formulated. In particular, one such proposed resolution (Zwanzig et al., 1992) can be shown unequivocally not to resolve the computational complexity form of the paradox, and in related arguments (Karplus and Shakhnovich, 1992) has been shown to lead to physically incorrect consequences (Section 5.3). For the paradox to be meaningful, it must be "falsifiable"—it must be possible to know when the paradox has been resolved.

In addition to restricted forms of PSP, it would be useful to know the computational complexity of other tasks in structure prediction that appear easier than the general problem, but whose complexities are none the less uncertain.

The task of computing side-chain conformations given full knowledge of a protein's backbone conformation is one such problem. Case studies using simulated annealing (Lee and Subbiah, 1991) have suggested that packing effects may suffice to determine, in part, the side-chain conformations in a protein's core. The computational complexity of this packing problem is unknown. Because only short-range effects are present, the graph of possible side-chain–side-chain interactions can be known in advance, is sparse, and consists of vertices of low degree. Previous experience—for instance, with Ising spin-glass models (Barahona, 1982), graph colorability (Garey and Johnson, 1979, p. 191) and cartographic labeling (Formann and Wagner, 1991; Marks and Shieber, 1991)—illustrates that such neighborhood interactions can, on their own, give rise to NP-hardness. On the other hand, many problems that contain such neighborhood interactions are tractable if restrictions can be placed on the nature of the graph (Garey and Johnson, 1979), suggesting that the problem of finding a mutually acceptable set of side-chain conformations for a protein could be tractable. (One currently known algorithm for predicting side-chain conformations based on backbone positions achieves 70% to 80% accuracy for χ_1 and χ_2 angles [Dunbrack and Karplus, 1993].) Not knowing the computational complexity of side-chain structure prediction leaves the algorithm developer in the quandary of not knowing whether inexact methods are truly necessary, given the possible existence of a superior exact algorithm.

Acknowledgments. We thank Ron Unger for answering detailed questions and providing a preprint (Unger and Moult, 1993). Aviezri Fraenkel also kindly provided a preprint (Fraenkel, 1993). We thank Harry Lewis, Eugene Shakhnovich, and Jim Clark for reading and commenting on the manuscript. JTN is grateful for a Graduate Fellowship from the Fannie and John Hertz Foundation. This research was supported in part by grants from the National Science Foundation and the National Institutes of Health.

NOTES

[1] The Thermodynamic Hypothesis states that a protein's native fold is the configuration of globally minimal free energy. However, it is generally assumed that a protein's states of lowest free energy are similar enough in entropy to justify the use of potential energies instead of free energies as a computational convenience; potential energies are much faster and more straightforward to compute.

[2] For example, if only nonbonded interactions between nearest neighbors along the chain are significant, the global minimum structure can be predicted efficiently (Finkelstein and Reva, 1992).

[3] The term *combinatorial optimization* is normally reserved for problems in which the solution space is discrete. Throughout this chapter we use the term to refer

Biochemistry

© Copyright 1990 by the American Chemical Society Volume 29, Number 37 September 18, 1990

Perspectives in Biochemistry

Additivity of Mutational Effects in Proteins

James A. Wells

Protein Engineering Department, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080

Received April 19, 1990; Revised Manuscript Received May 29, 1990

The energetics of virtually all binding functions in proteins is the culmination of a set of molecular interactions. For example, removal of a single molecular contact by a point mutation causes relatively small reductions (typically 0.5–5 kcal/mol) in the free energy of transition-state stabilization [for reviews see Fersht (1987) and Wells and Estell (1988)], protein–protein interactions (Laskowski et al., 1983, 1989; Ackers & Smith, 1985), or protein stability [for review see Matthews (1987)] compared to the overall free energy associated with these functional properties (usually 5–20 kcal/mol). Thus, it is possible to modulate protein function by mutation at many contact sites. In fact, to design large changes in function will often require mutation of more than one functional residue.

There is now a large data base for free energy changes that result when single mutants are combined. A review of these data shows that, in the majority of cases, the sum of the free energy changes derived from the single mutations is nearly equal to the free energy change measured in the multiple mutant. However, there are two major exceptions where such simple additivity breaks down. The first is where the mutated residues interact with each other, by direct contact or indirectly through electrostatic interactions or structural perturbations, so that they no longer behave independently. The second is where the mutation causes a change in mechanism or rate-limiting step of the reaction. It is important to note that the additive effects discussed here do not change the molecularity of their respective reactions. When the molecularity of the reaction changes [as in comparing the free energy of binding of one linked substrate (A–B) versus the sum of two fragments (A plus B)], large deviations from simple additivity can result from entropic effects (Jencks, 1981). Although the focus here is on enzyme activity, similar conclusions may be drawn from mutations affecting protein–protein interactions, protein–DNA recognition, or protein stability. Some practical examples and applications are discussed.

Additivity Relationships

The change in free energy of a functional property caused by a mutation at site X is typically expressed relative to that of the wild-type protein as $\Delta\Delta G_{(X)}$. Such free energy changes for two single mutants (X and Y) can be related to those of a double mutant (designated X,Y) by eq 1 (Carter et al., 1984; Ackers & Smith, 1985). The ΔG_I term (also called the coupling energy; Carter et al., 1984) should reflect the extent to which the change in energy of interaction between sites X and Y affects the functional property measured. It is possible for ΔG_I to be either positive or negative depending upon whether the interactions between the mutant side chains reduce or enhance the functional property measured. Furthermore, the ΔG_I term should not exceed the free energy of interaction between side chains at sites X and Y except in cases where these mutations cause large structural perturbations. This was first applied to evaluating the functional independence of residues mutated in tyrosyl-tRNA synthetase (Carter et al., 1984). In one case the sum of the $\Delta\Delta G$ values for single mutants was equal to that of the double mutant, indicating the sites functioned independently; in another example there was a large discrepancy, suggesting the sites were interacting.

$$\Delta\Delta G_{(X,Y)} = \Delta\Delta G_{(X)} + \Delta\Delta G_{(Y)} + \Delta G_I \qquad (1)$$

Simple Additivity in Transition-State Binding Interactions

The strengths of noncovalent interactions are strongly dependent upon the nature of the two groups and the distance (r) between them. For example, the free energy of charge–charge, random charge–dipole, random dipole–dipole, van der Waals attraction, and repulsion decay as $1/r$, $1/r^4$, $1/r^6$, $1/r^6$, and $1/r^{12}$, respectively [for review see Fersht (1985)]. Thus, when the side chains at sites X and Y are remote to one another and assuming no large structural perturbations, the ΔG_I term should be negligible and eq 1 thus simplifies to

$$\Delta\Delta G_{(X,Y)} \simeq \Delta\Delta G_{(X)} + \Delta\Delta G_{(Y)} \qquad (2)$$

This situation, here referred to as simple additivity, is generally observed except where side chains are close to each other or when one or both of the mutants change the rate-limiting step or reaction mechanism. These principles are well illustrated from data of additive mutational effects on transition-state stabilization energies.

0006-2960/90/0429-8509$02.50/0

8509-8517

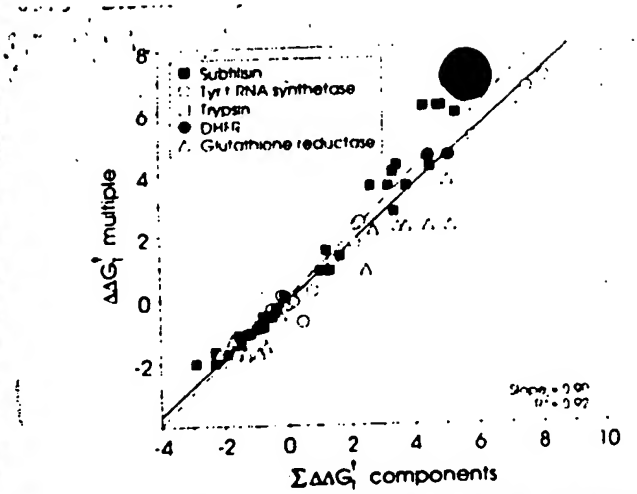

FIGURE 1: Plot of the changes in transition-state stabilization energies for the multiple mutant versus the sum for the component mutants. Data are taken from Table I and represent mutants from subtilisin (■), tyrosyl-tRNA synthetase (○), trypsin (□), DHFR (●), and glutathione reductase (△), where mutant or wild-type side chains should not contact one another. The dashed line has a slope of 1, and the solid line is a best fit to all the data.

Changes in transition-state stabilization energy ($\Delta\Delta G_T^\circ$) caused by a mutation can be calculated from eq 3 (Wilkinson et al., 1983), in which R is the gas constant, T is the absolute

$$\Delta\Delta G_T^\circ = -RT \ln \frac{(k_{cat}/K_M)_{mutant}}{(k_{cat}/K_M)_{wild\text{-}type}} \qquad (3)$$

temperature, k_{cat} is the turnover number, and K_M is the Michaelis constant for the mutant and wild-type enzyme against a fixed substrate. $\Delta\Delta G_T^\circ$ represents the change in free energy to reach the transition-state complex (E·S°) from the free enzyme and substrate (E + S).

To analyze the proposition that the interaction energy term, $\Delta G_{T(I)}^\circ$, is relatively small when the sites of mutation (X and Y) are remote to one another, $\Delta\Delta G_T^\circ$ values were collected from the literature where side-chain substitutions in the multiple mutant are beyond van der Waals contact (>4 Å distant) from each other (Table I). There are at least 25 examples distributed across five different enzymes where $\Delta\Delta G_T^\circ$ values can be calculated for the individual and multiple mutants assayed in at least two different ways. Among these are examples where electrostatic interactions, hydrogen bonding, and steric and hydrophobic effects have been altered separately or in combination with others. The X-ray structures of the wild-type proteins show that the wild-type side chains are not in contact. Modeling suggests the mutant side chains are beyond possible van der Waals contact unless the mutant side chains were to cause significant changes in the overall protein structure. Such large changes are rarely observed in structures of site-specific mutant proteins (Katz & Kossiakoff, 1986; Alber et al., 1987; Howell et al., 1986; Wilde et al., 1988) or even highly variant natural proteins (Chothia & Lesk, 1986).

A collective plot of the sum of the $\Delta\Delta G_T^\circ$ values for the component mutants versus the corresponding multiple mutant (Table I) gives a remarkably strong correlation ($R^2 = 0.92$) with a slope near unity (Figure 1). The simplest interpretation is that the interaction term, $\Delta G_{T(I)}^\circ$, is small compared to the overall effects on $\Delta\Delta G_{T(X,Y)}^\circ$. It is formally possible that there are large and compensating effects between side chains X and Y that systematically lead to small net values for $\Delta G_{T(I)}^\circ$.

There are some notable exceptions that weaken the correlation within the data set (Table I). In particular, combining the R204L mutation in *Escherichia coli* glutathione reductase gives a less than additive effect, especially when combined with another mutant, R198M (Scrutton et al., 1990). These basic residues are not in direct contact, but both side chains form a salt bridge with the 2′-phosphate group of NADPH. Indeed, the largest discrepancies are when these mutants are assayed with NADPH as compared to NADH. Similarly, the sum of the $\Delta\Delta G_T^\circ$ values for two positively charged component mutants in subtilisin (D99K and E156K) overestimates the effect of the multiple mutant when assayed with an Arg but not with a Phe substrate (Russell & Fersht, 1987). Such discrepancies are not too surprising because charge–charge interactions fall off as $1/r$ and can exhibit long-range effects in proteins [for example, see Russell and Fersht (1988)]. The physical basis for other large discrepancies not involving electrostatic substitutions is less clear but may involve unexpectedly large structural changes or changes in enzyme mechanism (see below).

These additivity tests are not particularly dominated by one of the single mutants in the sum. The average contribution (±SE) for the most dominant mutant in ech sum calculated from the 69 additivity tests given in Table I is only 68% (±15%) of the total sum (theoretical is ~50%). Furthermore, the plot in Figure 1 is not analogous to graphs of correlated variables, where A is plotted versus the sum of A + B, because in Figure 1 the values on the y-axis are determined *independently* from those on the x-axis.

Complex Additivity in Transition-State Stabilization—When $\Delta G_{T(I)}^\circ \neq 0$

(A) Change in Interaction Energy between Sites X and Y. Where residues X and Y are close enough to contact, it is more likely that the $\Delta G_{T(I)}^\circ$ term will be significant. There are 11 examples collectively from tyrosyl-tRNA synthetase and subtilisin that fit this category (Table II).

A series of mutants in tyrosyl-tRNA synthetase at positions 48 and 51 (Carter et al., 1984; Lowe et al., 1985) show complex additivity (Table II). His48 and Thr51 in the wild-type structure are next to each other on adjacent turns of an α-helix. His48 hydrogen bonds to the ribose ring oxygen of ATP while Thr51 can make van der Waals contact with ATP. The T51P mutation increases the catalytic efficiency of the enzyme in some assays by more than −2 kcal/mol (Wilkinson et al., 1984). However, when this mutation is combined with mutations at position 48, the effects are not simply additive. An X-ray structure of the T51P mutant indicates there are no structural changes in the α-helix (Brown et al., 1987). Instead, it is suggested that the T51P mutant is improved over wild type because the wild-type enzyme contains a bound water in the vicinity of Thr51 that disfavors substrate binding. Blow and co-workers (Brown et al., 1987) argue that the change in solvent structure propagated to position 48 may account for the complex additivity. In the previous section, the double mutant (H48G,T51A) exhibited nearly simple additivity (Table I). Presumably, the smaller and less hydrophobic alanine substitution at position 51 should not introduce as large a change in solvent structure as the pyrrolidone ring of proline.

In the case of subtilisin (Table II), Glu156 is near the top of the P1 binding crevice while Gly166 is at the bottom. In the wild-type enzyme these sites do not make direct van der Waals contact, but large side chains substituted at position 166 can be modeled to contact the residue at position 156. In fact, X-ray structural analysis shows that an Asn side chain at position 166 makes a good hydrogen bond with Glu156 (Bott et al., 1987). Moreover, all of the substitutions are polar or charged, the energetics of which are expected to be the most long range. Thus, the mutant side chains alter substantially the intramolecular interactions between positions 156 and 166.

Table I: Comparison of Sums of $\Delta\Delta G_T^\circ$ from Component Mutants vs the Multiple Mutant Where the Mutant or Wild-Type Side Chains *Do Not* Contact One Another

assay	$\Delta\Delta G_T^\circ$ component mutants			sum	multiple mutant
Tyrosyl-tRNA Synthetase					
	C35G + H48G[a]				
ATP/PP_i	+1.20	+1.04		+2.24	+2.30
ATP/tRNA	+1.05	+1.13		+2.18	+1.68
Tyr/PP_i	+1.14	+1.12		+2.26	+2.32
Tyr/tRNA	+0.32	+1.12		+1.45	+1.20
	C35G + T51P				
ATP/PP_i	+1.20	−1.91		−0.71	−1.14
ATP/tRNA	+1.05	−2.35		−1.30	−1.88
Tyr/PP_i	+1.14	−0.64		+0.50	−0.74
Tyr/tRNA	+0.32	+0.50		+0.82	+0.21
	C35G + T51C[b]				
ATP/tRNA	+1.05	−0.93		+0.12	−0.22
ATP/Tyr	+1.14	−0.91		+0.23	−0.13
	H48N + T51A[c]				
ATP/PP_i	+0.26	−0.38		−0.12	+0.04
ATP/tRNA	−0.13	−0.32		−0.45	−0.37
	T40A + H45G[d]				
Tyr/Tyr	+5.02	+3.15		+8.17	+6.95
ATP/Tyr	+5.13	+2.44		+7.57	+6.67
Rat Trypsin					
	G216A + G226A[e]				
K	+2.75	+3.13		+5.88	+5.07
R	+2.19	+4.91		+7.10	+5.90
Dihydrofolate Reductase ($\Delta\Delta G_{binding}$)					
	F31V + L54G[f]				
H_2F	+1.6	+2.9		+4.5	+4.5
MTX	+2.2	+2.9		+5.1	+4.5
Subtilisin BPN′					
	E156S + Y217L + G169A[g]				
E	−1.43	−0.87	−0.62	−2.92	−2.06
Q	−0.60	−0.36	−0.32	−1.28	−1.14
A	−0.15	−0.41	−0.27	−0.83	−0.92
K	+1.70	−0.08	−0.30	+1.32	+0.87
M	−0.86	−0.32	−0.39	−1.57	−1.41
F	−0.61	−0.29	−0.66	−1.56	−1.17
Y	−0.24	−0.12	−0.41	−0.77	−0.59
	E156S + Y217L				
E	−1.43	−0.87		−2.30	−1.67
Q	−0.60	−0.36		−0.96	−0.96
A	−0.15	−0.41		−0.56	−0.53
K	+1.70	−0.08		+1.62	+1.33
M	−0.86	−0.32		−1.18	−1.11
F	−0.61	−0.29		−0.90	−0.84
Y	−0.24	−0.12		−0.36	−0.32
	E156S, Y217L + G169A				
E	−1.67	−0.62		−2.29	−2.06
Q	−0.96	−0.32		−1.28	−1.14
A	−0.53	−0.27		−0.80	−0.92
K	+1.33	−0.30		+1.03	+0.87
M	−1.11	−0.39		−1.50	−1.41
F	−0.84	−0.66		−1.50	−1.17
Y	−0.32	−0.41		−0.73	−0.59
	D99S + E156S[h]				
R	+0.47	+0.77		+1.24	+1.52
F	0	−0.62		−0.62	−0.52
Subtilisin BPN′					
	D99K + E156K				
R	+1.29	+2.12		+3.41	+2.74
F	+0.13	−0.49		−0.36	−0.42
	E156S, G166A + G169A, Y217L[i]				
F	−0.40	−1.46		−1.86	−1.76
Y	+0.94	−1.03		−0.09	+0.02
	G166A + S24C, H64A				
F	−0.40	+4.96		+4.56	+4.11
Y	+0.94	+4.40		+5.34	+5.84
	E156S, G169A, Y217L + S24C, H64A				
F	−1.46	+4.96		+3.50	+4.21
Y	−1.03	+4.40		+3.37	+3.96
	S24C, H64A, G169A, Y217L + G166A				
F	+4.21	−0.40		+3.81	+3.53
Y	+3.96	+0.94		+4.90	+6.07
	S24C, H64A, G166A + E156S, G169A, Y217L				
F	+4.11	−1.46		+2.65	+3.53
Y	+5.84	−1.03		+4.81	+6.07
	S24C, H64A + E156S, G166A, G169A, Y217L				
F	+4.96	−1.76		+3.20	+3.53
Y	+4.40	+0.02		+4.38	+6.07
***E. coli* Glutathione Reductase**					
	A179G + R198M[j]				
NADH	−1.10	−0.62		−1.72	−1.32
NADPH	+0.08	+2.68		+2.76	+2.11
	A179G + R204L				
NADH	−1.10	+0.41		−0.69	−1.54
NADPH	+0.08	+2.42		+2.50	+0.87
	R198M + R204L				
NADH	−0.62	+0.41		−0.21	−0.51
NADPH	+2.68	+2.42		+5.10	+3.70
	A179G + R179M, R204L				
NADH	−1.10	−0.51		−1.61	−1.72
NADPH	+0.08	+3.70		+3.78	+2.22
	R198M + A179G, R204L				
NADH	−0.62	−1.54		−2.16	−1.72
NADPH	+2.68	+0.87		+3.55	+2.22
	R204L + A179G, R198M				
NADH	+0.41	−1.32		−0.91	−1.72
NADPH	+2.42	+2.11		+4.53	+2.22
	R179G + R198M + R204L				
NADH	−1.10	−0.62	+0.41	−1.31	−1.72
NADPH	+0.08	+2.68	+2.42	+5.18	+2.22

[a] Carter et al. (1984). The assays refer to measurements of ATP-dependent pyrophosphate exchange (ATP/PP_i) or tRNA charging (ATP/tRNA) under saturating conditions for tyrosine and vice versa for Tyr/PP_i exchange and Tyr/tRNA charging. [b] Lowe et al. (1985). The ATP/Tyr activation assay refers to formation of tyrosyl adenylate under saturating concentrations of tyrosine. [c] Jones et al. (1986). [d] Leatherbarrow et al. (1986). The ATP/Tyr and Tyr/Tyr activation assays refer to formation of tyrosyl adenylate under pre-steady-state conditions, and k_{cat}/K_M is calculated from k_3/K_S for tyrosine and ATP, respectively. [e] Craik et al. (1985). The substrate was D-Val-Leu-(X)-aminofluorocoumarin where the P1 residue (X) is either Lys (K) or Arg (R). [f] Mayer et al. (1986). The ligand was either dihydrofolate (H_2F) or methotrexate (MTX). [g] Wells et al. (1987a). The substrate was succinyl-L-Ala-L-Ala-L-Pro-L-(X)-*p*-nitroanilide where the P1 (X) residue (Schechter & Berger, 1937) was either Glu (E), Gln (Q), Ala (A), Lys (K), Met (M), Phe (F), or Tyr (Y). [h] Russell and Fersht (1987). The substrate was benzoyl-L-Val-Gly-L-Arg-*p*-nitroanilide (R) or succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide (F). [i] Carter et al. (1989). The substrate was succinyl-L-Phe-L-Ala-L-His-L-(X)-*p*-nitroanilide where X was either Phe (F) or Tyr (Y). [j] Scrutton et al. (1990). The assay followed the reduction of oxidized glutathione by NADH or NADPH.

Table II: Comparison of Sums of $\Delta\Delta G_T^\circ$ [redacted] Component Mutants vs the Multiple Mutant Where the Mutant Side Chains *Can* Contact One Another.

assay[a]	$\Delta\Delta G_T^\circ$ component mutants		$\Delta\Delta G_T^\circ$ sum	$\Delta\Delta G_T^\circ$ multiple mutant
Tyrosyl-tRNA Synthetase				
	H48G + T51P[b]			
ATP/PP_i	+1.04	−1.91	−0.87	+1.07
ATP/tRNA	+1.13	−2.35	−1.22	+0.77
Tyr/PP_i	+1.12	−0.64	+0.48	+1.02
Tyr/tRNA	+1.12	+0.50	+1.63	+0.17
ATP/Tyr	+0.95	−1.99	−1.04	+1.04
Tyr/ATP	+1.07	−0.38	+0.69	+0.82
	H48N + T51P			
ATP/Tyr	+0.18	−1.99	−1.81	−0.76
Tyr/Tyr	+0.36	−0.38	−0.02	−0.64
ATP/tRNA	−0.02	−2.23	−2.25	−1.07
	N48G + T51P			
ATP/Tyr	+0.37	−0.94	−0.57	+0.86
Tyr/Tyr	+0.41	−1.00	−0.59	+0.45
ATP/tRNA	+1.26	−1.05	+0.21	+0.90
	Q48G + T51P			
ATP/Tyr	−1.31	−1.09	−2.40	−1.22
Tyr/Tyr	−2.05	−1.65	−3.70	−2.31
ATP/tRNA	−1.87	−1.85	−3.72	−2.23
	H48Q + T51P			
ATP/Tyr	+2.26	−1.99	+0.27	+1.17
Tyr/Tyr	+3.13	−0.38	+2.75	+1.48
ATP/tRNA	+3.11	−2.23	+0.88	+1.26
Subtilisin BPN'				
	E156Q + G166D[d]			
Q	−1.04	+1.27	+0.23	+0.75
M	−0.45	+1.83	+1.38	+0.16
K	+2.15	+0.53	+2.68	+0.26
	E156S + G166D			
Q	−0.59	+1.27	+0.68	+0.74
M	−0.85	+1.83	+0.98	+0.66
K	+1.68	+0.53	+2.22	+0.49
	E156Q + G166N			
E	−1.71	−0.11	−1.82	−0.69
Q	−1.04	+0.14	−0.90	−0.77
M	−0.45	+0.18	−0.27	−1.10
K	+2.15	+0.48	+2.73	+1.16
	E156S + G166N			
E	−1.44	−0.11	−1.55	−0.51
Q	−0.59	+0.14	−0.45	−0.85
M	−0.85	+0.18	−0.67	−0.78
K	+1.68	+0.48	+2.16	+1.26
	E156S + G166K			
E	−1.44	−3.49	−4.93	−4.49
Q	−0.59	−1.03	−1.62	−0.95
M	−0.85	−1.37	−2.22	−1.12
K	+1.68	+0.51	+2.19	+1.88
	E156Q + G166K			
E	−1.71	−3.49	−5.20	−4.49
Q	−1.04	−1.03	−2.07	−0.95
M	−0.45	−1.37	−1.82	−1.12
K	+2.15	+0.51	+2.66	+1.88

[a] See Table I for description assays. [b] Lowe et al. (1985). [c] Carter et al. (1984). [d] Wells et al. (1987b).

In these six examples there are large and systematic discrepancies between the sum of the $\Delta\Delta G_T^\circ$ values for the single mutants and those of the corresponding double mutant (Wells et al., 1987b). In almost all cases, the sum of the $\Delta\Delta G_T^\circ$ values for the single mutants is much greater than the value for the multiple mutant. Nonetheless, the $\Delta\Delta G_T^\circ$ value predicted from the sum of the single mutants does have the same sign as that for the double mutant, so that the single mutants predict qualitatively the effect on the multiple mutant.

A plot (Figure 2) of the collective data set from Table II is in contrast to that seen in Figure 1. The $\Delta\Delta G_T^\circ$ values for the multiple mutants correlate more poorly with the sum of

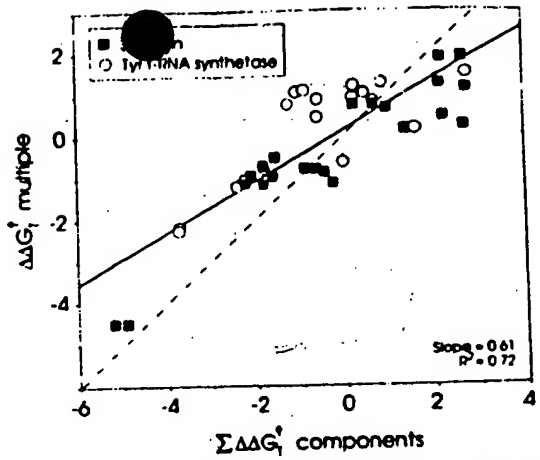

FIGURE 2: Data are taken from Table II for mutants of subtilisin (■) or tyrosyl-tRNA synthetase (○) where mutant or wild-type side chains can contact each other. The dashed line represents a theoretical line of unity slope, and the solid line represents the best fit.

the component single mutants ($R^2 = 0.72$). Moreover, the slope of the line (0.61) is much below unity. This indicates that the function of one residue is compromised by mutation of another. Of the 40 additivity examples, the average contribution of the most dominant single mutant to the sum of the $\Delta\Delta G_T^\circ$ values is 71% (±13%) of the total. Thus (as in Figure 1), both single mutants can contribute substantially to free energy changes measured in the multiple mutant. However, this data set is derived from mutations at only two different sites on two different proteins.

In summary, complex additivity can be observed when mutations at sites X and Y change the intramolecular interaction energy between sites. This can be mediated by direct steric, electrostatic, hydrogen-bonding, or hydrophobic interactions or indirectly through large structural changes in the protein, solvent shell, or electrostatic interactions. Complex additivity is most likely to occur where the sites of mutation are very close together and larger or chemically divergent side chains are introduced.

(B) Mutations at Sites X or Y Change the Enzyme Mechanism or Rate-Limiting Step. If the catalytic functions of two or more residues are interdependent, then a mutation of one residue can affect the functioning of the other(s). This form of complex additivity is well illustrated for mutations in the catalytic triad and oxyanion binding site of subtilisin (Carter & Wells, 1988, 1990). In the catalytic mechanism of subtilisin (Figure 3), the rate-limiting step in amide bond hydrolysis is transfer of the proton from Ser221 to His64 with nucleophilic attack upon the scissile carbonyl carbon. This is accompanied by electrostatic stabilization of the protonated imidazole by Asp32 and hydrogen bonding to the oxyanion by the side chain of Asn155 and the main-chain amide of Ser221. Mutational analysis shows that once the catalytic Ser221 is mutated to Ala (S221A), additional mutations in the triad or oxyanion binding site cause no further loss in catalytic efficiency (Table III).

The S221A enzyme retains a catalytic activity that is still 10^4 above the solution hydrolysis rate (Carter & Wells, 1988). It is proposed that this residual activity is derived from remaining transition-state binding contacts outside of the catalytic triad coupled with solvent attack upon the carbonyl carbon from the face opposite position 221 (Carter & Wells, 1990). This proposal is based on a model showing that there is no room for a water molecule near Ala221 once the substrate is bound. Furthermore, conversion of Asn155 to Gly enhances the activity of the S221A mutant by −1.2 kcal/mol (Table III).

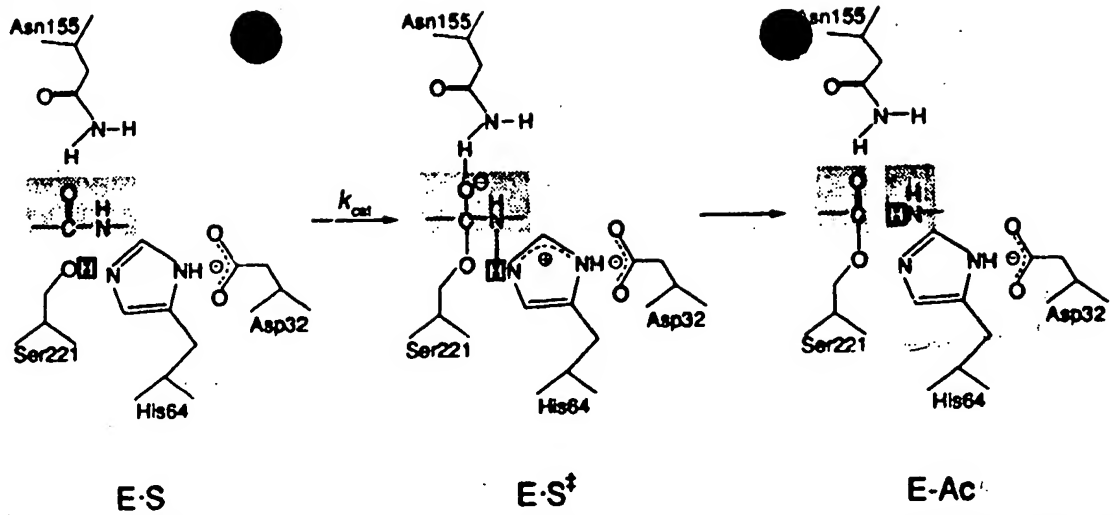

FIGURE 3: Schematic diagram of the mechanism of subtilisin showing the rate-limiting acylation step for hydrolysis of peptide bonds. Reproduced with permission from Carter and Wells (1988). Copyright 1988 Macmillan.

Table III: Comparison of Sums of $\Delta\Delta G_T^\circ$ from Component Mutants vs the $\Delta\Delta G_T^\circ$ for Multiple Mutants in the Catalytic Triad and Oxyanion Binding Site of Subtilisin BPN'[a]

component mutants			sum	multiple mutant
S221A + H64A[b]				
+8.93	+8.84		+17.76	+8.83
S221A + D32A				
+8.93	+6.52		+15.45	+8.86
H64A + D32A				
+8.84	+6.52		+15.36	+7.48
S221A + H64A + D32A				
+8.93	+8.84	+6.52	+24.29	+8.65
S221A + H64A, D32A				
+8.93	+7.48		+16.40	+8.65
H64A + S221A, D32A				
+8.84	+8.86		+17.70	+8.65
D32A + S221A, H64A				
+6.52	+8.83		+15.35	+8.65
S221A + N155G[c]				
+8.93	+3.08		+12.01	+7.70

[a] All enzymes were assayed with the substrate succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide. [b] Carter and Wells (1988). [c] Carter and Wells (1990).

This is consistent with the opposite-face solvent attack mechanism of S221A, because the oxyanion (Figure 3) would develop away from Asn155 and the N155G mutation improves solvent accessibility to the scissile carbonyl carbon.

Complex additivity is also seen for subtilisin mutated at positions 64 and 32. The double (H64A,D32A) and corresponding single mutants show a linear dependence upon hydroxide ion concentration (between pH 8 and 10) that may reflect hydroxide assistance in the deprotonation of the Oγ of Ser221 (Carter & Wells, 1988). Thus, once His64 is converted to Ala, Asp32 is a liability, presumably by electrostatic repulsion of hydroxide ion. [Note the −1.3 kcal/mol improvement in $\Delta\Delta G_T^\circ$ for the double mutant (H64A,D32A) compared to H64A alone; Table III.]

In summary, if an enzyme mechanism relies upon cooperative interaction between two or more residues, then multiple mutations within this subset can result in large values for $\Delta G_{T(I)}^\circ$. In fact, if the mechanism is changed substantially, residues that were a catalytic asset can become a liability. Simple additivity can also break down when one or more of the mutations cause a change in the rate-limiting step. In an extreme case, one may have a number of mutants in an enzyme that enhance the activity, but the cumulative enhancement of activity could not go beyond the diffusion-controlled limit (Albery & Knowles, 1976).

Additive Effects on Substrate Binding

The analysis above considered changes in binding free energies between the free enzyme and substrate (E + S) to yield the bound transition-state complex (E·S‡). The steady-state kinetic analysis for subtilisin and tyrosyl-tRNA synthetase is such that the K_M values approximate the enzyme–substrate dissociation constant K_s. Additivity analysis based on calculations of $\Delta\Delta G_{binding}$ (from K_M values) or $\Delta\Delta G_{cat}$ (from k_{cat} values) yields qualitatively the same results (not shown) as shown in Tables I and II and Figures 1 and 2. Thus, deviations from simple additivity are not systematically found in either the energetics to form the E·S complex or those to reach E·S‡.

Additive Effects on Protein–Protein Interactions

The first clear examples of additive binding effects caused by amino acid replacements in proteins were reported by Laskowski et al. (1983) and reviewed by others (Ackers & Smith, 1985; Horovitz & Rigbi, 1985). One hundred natural variants of a proteinase inhibitor, the ovomucoid third domain, have been isolated and sequenced from the eggs of different bird species (Empie & Laskowski, 1982; Laskowski et al., 1987). This is a nested set of proteins because for any one of these avian inhibitors there is a close relative containing only one or a few amino acid substitutions. Moreover, the association constants (K_a) of these inhibitors with a variety of serine proteinases vary over an enormous range (10^9-fold). Laskowski et al. (1983, 1989) have shown that the effect of a given residue replacement on K_a is about the same irrespective of the inhibitor scaffold the replacement is made in.

In addition to ovomucoid, four additivity examples have been constructed from natural variants at the subunit interface of tetrameric hemoglobin (Ackers & Smith, 1985). Three additivity examples have been analyzed for interactions of hGH with its receptor (B. C. Cunningham and J. A. Wells, unpublished results) and one example for association of synthetic variants of the RNase S peptide with RNase S protein (Mitchinson & Baldwin, 1986). The entirety of this data set is not tabulated because much on the ovomucoid inhibitors and hGH is unpublished. Nonetheless, these researchers were kind enough to provide their data formatted so it could be plotted collectively in Figure 4. These data consist of 91 additivity examples (80 in ovomucoids alone), representing 22 multiple mutants across four different proteins, and span a wide range of change in binding free energy (−10 to +7

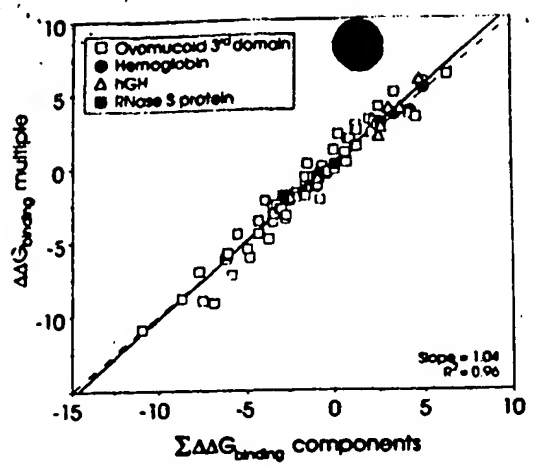

FIGURE 4: Plot showing the sum of changes in free energies of binding at protein–protein interfaces for component mutants versus the corresponding multiple mutant. Data represent interactions between ovomucoid third domain and various serine proteases (□) (R. Wynn and M. Laskowski, personal communication), regulatory interface of $\alpha_2\beta_2$ hemoglobin (●) (Ackers & Smith, 1985), hGH and its receptor (stippled △) (B. Cunningham and J. Wells, personal communication), and RNase S peptide and S protein (■) (Mitchinson & Baldwin, 1986). The dashed line represents a line of unity slope, and the solid line is the best fit.

kcal/mol). The plot shows a very strong linear correlation (R^2 = 0.96) with a slope near unity. Although the data for the ovomucoid were not sorted to evaluate changes at intramolecular contact sites, most are not expected to be in contact, and all of the other examples represent noncontact sites. Thus, the large data base derived from natural variants of ovomucoid third domain, as well as a smaller number of examples from several other proteins, indicates that multiple mutations at protein–protein interfaces commonly produce simple additive effects.

Additive Effects in DNA–Protein Interactions

One of the clear advantages in analyzing DNA–protein interactions is the ability to apply powerful selections that make analysis by random mutational studies feasible. Additivity in DNA–protein interactions was first demonstrated by reversion analysis of λ repressor (Nelson & Sauer, 1985). A mutation that decreased the binding affinity for the λ operator site (K4Q) was reverted by mutations at several second sites (E34K, G48S, and E83K). When these second-site revertants were introduced into wild-type λ repressor, they caused increases in affinity similar to those observed in the first-site suppressor mutant (K4Q).

Functional independence for mutations at DNA–protein contacts has been demonstrated by additive effects for mutants of CAP (catabolite gene activator protein) and its operator sequence (Ebright et al., 1987) as well as *lac* repressor and its corresponding operator sequence (Ebright, 1986). Simple additivity of mutational effects in the operator sequences for Cro repressor (Takeda et al., 1989) and λ repressor (Sarai & Takeda, 1989) has been most systematically demonstrated. Simple additivity has also been reported for multiple mutations in the *lac* repressor (Lehming et al., 1990). In fact, simple additivity is so predictable in DNA–protein interactions that the observation of complex additivity has been used to predict specific DNA–protein contacts in the *lac* repressor–operator complex (Ebright, 1986).

Additive Effects on Protein Stability

The first systematic analysis of additive effects of site-specific mutations on protein stability was reported by Shortle and Meeker (1986). Five multiple mutants in staphylococcal nuclease were constructed from a group of random single mutants that were screened initially for their ability to affect the stability of the enzyme in vivo. The component mutants do not make direct contact with each other in the multiple mutants. Generally, these variants exhibit nearly additive effects except for the double mutant V66L,G88V (Table IV). In addition to those of staphylococcal nuclease, additive effects on the $\Delta\Delta G_{unfolding}$ (assayed by reversible denaturation) have also been determined for the N-terminal domain of λ repressor (one example; Hecht et al., 1986), the α-subunit of *E. coli* Trp synthetase (one example; Hurle et al., 1986), T4 lysozyme (six examples; Wetzel et al., 1988), the gene V product of bacteriophage f1 (one example; Sandberg & Terwilliger, 1989),

Table IV: Comparison of Sums of $\Delta\Delta G_{unfolding}$ from Component Mutants vs the Multiple Mutant

assay	$\Delta\Delta G_{unfolding}$ component mutants			sum	multiple mutant
Staphylococcal Nuclease					
	V66L + G79S[a]				
GuHCl	−0.2	−2.6		−2.8	−3.3
urea	+0.2	−2.9		−2.7	−3.6
	V66L + G88V				
GuHCl	−0.2	−1.0		−1.2	−2.1
urea	+0.2	−0.9		−0.7	−1.4
	I18M + A69T				
GuHCl	−0.6	−2.7		−3.3	−2.8
urea	−0.7	−2.9		−3.6	−3.8
	I18M + A90S				
GuHCl	−0.6	−1.4		−2.0	−2.2
urea	−0.7	−1.4		−2.1	−2.2
	V66L + G79S + G88V				
GuHCl	−0.2	−2.6	−1.0	−3.8	−3.0
urea	+0.2	−2.9	−0.9	−3.6	−3.4
N-Terminal Domain of λ Repressor					
	G46A + G48A[b]				
thermal melt	+0.7	+0.9		+1.6	+1.1
T4 Lysozyme					
	I3C + C54V[c]				
thermal melt	+1.2	−0.7		+0.5	+0.4
	I3C + C54T				
thermal melt	+1.2	+0.3		+1.5	+1.5
	I3C + C54T + R96H				
thermal melt	+1.2	+0.3	−2.8	−1.3	−2.5
	I3C,C54T + R96H				
thermal melt	+1.5	−2.8		−1.3	−2.5
	I3C + C54T + A146T				
thermal melt	+1.2	+0.3	−1.5	0	−0.5
	I3C,C54T + A146T				
thermal melt	+1.5	−1.5		0	−0.5
Bacteriophage f1 Gene V					
	V35I + I47V[d]				
GuHCl	−0.4	−2.4		−2.8	−2.9
Kringle-2 of tPA					
	H64Y + R68G[e]				
thermal melt	+2.9	+0.7		+3.6	+3.4
Turkey Ovomucoid Third Domain					
	G32A + N28S[f]				
thermal melt	+0.8	−0.5		+0.3	+0.2
	Y20H + N45-CHO				
thermal melt	−0.8	+0.3		−0.5	−0.6
α Subunit of *E. coli* Trp Synthetase					
	Y175C + G211E[g]				
GuHCl	−0.1	+0.3		+0.2	−1.3

[a] Shortle and Meeker (1986). [b] Hecht et al. (1986). [c] Wetzel et al. (1988). [d] Sandberg and Terwilliger (1989). [e] R. Kelley, personal communication. [f] Otlewski and Laskowski (1990). N45-CHO refers to a glycosylation of Asn45. [g] Hurle et al. (1986).

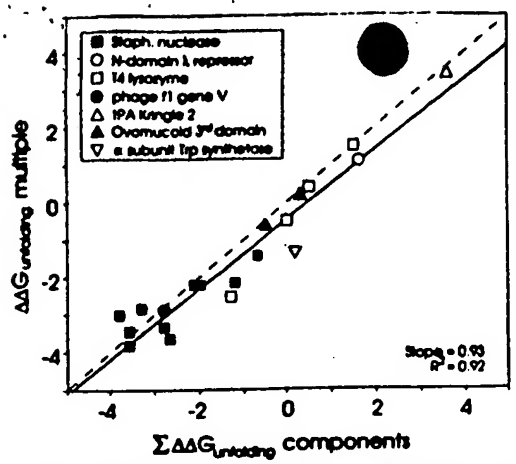

FIGURE 5: Plot showing sum of changes in free energy of unfolding of component mutants and resulting multiple mutant. Data are taken from Table IV and represent staphylococcal nuclease (■), N-terminal domain of λ repressor (○), T4 lysozyme (□), bacteriophage f1 gene V product (●), Kringle-2 domain of tissue plasminogen activator (▲), turkey ovomucoid third domain (△), and the α-subunit of Trp synthetase (▽). The dashed line represents a theoretical line of unity slope, and the solid line represents the best fit.

natural variants of ovomucoid third domain (two examples; Otlewski & Laskowski, 1990), and the Kringle-2 domain of human tissue plasminogen activator (t-PA) (one example; R. Kelley, personal communication).

Collectively, this data set gives a high linear correlation (R^2 = 0.94) and slope near unity (Figure 5). The generally simple additive behavior is somewhat surprising given the highly cooperative nature of protein folding. There are discrepancies in some of the additivity examples besides the staphylococcal nuclease mutant (V66L,G88V). For example, the 1.5 kcal/mol discrepancy for the Y175C,G271E double mutant in Trp synthetase (Table IV) is proposed to result from the fact that these residues are in direct contact (Hurle et al., 1986). Furthermore, proximity effects may account for the large differences between the sum of the component mutants and the multiple mutants for the α-helical double glycine mutant G46A,G48A in λ repressor (Hecht et al., 1986), and when combining R96H with the C3–C97 disulfide mutant in T4 lysozyme (Wetzel et al., 1988). In contrast, an exchange of two side chains that contact one another (V35I and I47V) in the hydrophobic core of the gene V product of f1 phage produced simple additive effects (Sandberg & Terwilliger, 1989; Table IV). It should be noted that this data base exhibiting simple additivity may be biased for single mutants that stably fold, because severely unstable proteins are more difficult to express.

By analogy to transition-state binding effects, one can certainly imagine instances where the stabilizing effects of mutations should reach a plateau. For example, denaturation at high temperatures can become controlled by a chemical step such as deamidation (Ahern et al., 1987), so that additional mutants that stabilize the folded form of the protein may be irrelevant. Another obvious example where complex additivity can be observed in protein stability is the stabilizing effect of disulfide bonds and noncovalent intramolecular contacts that require interactions between two or more residues. In these cases, the stabilizing interaction between two side chains can be broken with only one mutation.

Applications of Additivity in Rational Protein Design

A strategy of additive mutagenesis, where a series of single mutants each making a small improvement in function are combined, is one of the most powerful tools in designing functional properties of proteins. This approach has been remarkably successful in stabilizing proteins to irreversible inactivation, such as λ repressor (Hecht et al., 1986), subtilisin (Bryan et al., 1987; Cunningham & Wells, 1987; Pantoliano et al., 1989), kanamycin nucleotidyltransferase (Liao et al., 1986; Matsumura, 1986), neutral protease (Imanaka et al., 1986), and T4 lysozyme (Wetzel et al., 1988; Matsumara et al., 1989). This strategy has been applied to enhancing the catalytic efficiency of a weakly active variant of subtilisin (Carter et al., 1989), engineering the substrate specificity of subtilisin (Wells et al., 1987a,b; Russell & Fersht, 1987) and the coenzyme specificity of glutathione reductase (Scrutton et al., 1990), designing protease inhibitors with exquisite protease specificity (Laskowski et al., 1989), and recruiting human prolactin to bind to the hGH receptor (Cunningham et al., 1990). In addition, additivity principles have been used to engineer the pH profile of subtilisin (Russell & Fersht, 1987) and to design the affinity and specificity of λ repressor (Nelson & Sauer, 1985).

For this approach to work does not require that all the component mutants act in a simply additive manner but just that their effects accomulate. For example, despite the complex additivity of effects in the catalytic triad of subtilisin, there are mutagenic pathways that are energetically cumulative for installing the triad (Carter & Wells, 1988; Wells et al., 1987c). Starting with the triple mutant S221A,H64A,D32A, there is a progressive enhancement for installing Ser221 (−1.1 kcal/mol), then His64 (−1.0 kcal/mol), and finally Asp32 (−6.5 kcal/mol). Another cumulative pathway of Ser221, then Asp32, and finally His64 is possible if the Ser221,Asp32 intermediate were to use HisP2 substrates (Carter & Wells, 1987). Elaborating such cumulative pathways is important for understanding how a catalytic apparatus may have evolved and is practically useful for considering how to install such catalytic machinery into weakly active catalytic antibodies.

Conclusions

In the majority of cases, combination of mutations that affect substrate or transition-state binding, protein–protein interactions, DNA–protein recognition, or protein stability exhibits simple additivity. Simple additivity is commonly observed for distant mutations at rigid molecular interfaces such as in protein–protein and DNA–protein interactions, where the mutations are unlikely to alter grossly the structure or mode of binding.

Large deviations from simple additivity can occur when the sites of mutations strongly interact with one another (by making direct contact or indirectly through electrostatic interactions or large structural perturbations) and/or when both sites function cooperatively (as for the catalytic triad and oxyanion binding site of subtilisin). Changes at sites that can contact each other do not always lead to complex additivity; this may reflect relatively weak interactions between the two sites or indicate that the interactions are compensatory and appear to be weak.

It is important to point out the magnitude of errors in predicting the free energy effect in the multiple mutant from the component single mutants. Generally, for those cases exhibiting simple additivity (Figures 1, 4, and 5), the discrepancy in free energy between the sums of the components and multiple mutants is about ±25%. Part of this is the result of compounding errors when summing the single mutants, and the rest is presumably due to weak interaction terms. Nonetheless, this means that if the total free energy change is about 3 kcal/mol, the change in the equilibrium constant

(related by $K_{xy}/K_{wt} = 10^{-\Delta\Delta G/RT} = 155$) will often be off by a factor of 4. Thus, while the free energy effects accumulate, significant deviations will occur in predicting the final equilibrium constants when component mutants contribute a large free energy term.

Simple additivity reflects the modularity of component amino acids in protein function. This results from the fact that the perturbations in energetics and structure resulting from most mutations are highly localized. In the past six years, an additive mutagenesis strategy has been extremely effective in engineering proteins—of course, nature has been using this strategy much longer.

Acknowledgments

I am grateful to Dr. Michael Laskowski and Rich Wynn for providing their data prior to publication on ovomucoid third domain and similarly to Brian Cunningham, Paul Carter, and Robert Kelley for making available their unpublished data. I am indebted for useful discussions with Drs. Michael Laskowski, Paul Carter, Jack Kirsch, and Tony Kossiakoff and many of my colleagues at Genentech and to Drs. Richard Ebright and William Jencks and those above for critical reading of the manuscript.

Registry No. RNase, 9001-99-4; tyrosyl-tRNA synthetase, 9023-45-4; trypsin, 9002-07-7; dihydrofolate reductase, 9002-03-3; subtilisin BPN', 9014-01-1; glutathione reductase, 9001-48-3; staphylococcal nuclease, 9013-53-0; lysozyme, 9001-63-2; plasminogen activator, 105913-11-9; tryptophan synthetase, 9014-52-2.

References

Ackers, G. K., & Smith, F. R. (1985) *Annu. Rev. Biochem. 54*, 597–629.

Ahern, T. J., Casal, J. I., Petsko, G. A., & Klibanov, A. M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 675–679.

Alber, T., Dao-pin, S., Wilson, K., Wozniak, J. A., Cook, S. P., & Matthews, B. W. (1987) *Nature 330*, 41–46.

Albery, W. J., & Knowles, J. R. (1976) *Biochemistry 15*, 5631–5640.

Ardelt, W., & Laskowski, M., Jr. (1990) *J. Mol. Biol.* (submitted for publication).

Bott, R., Ultsch, M., Wells, J., Powers, D., Burdick, D., Struble, M., Burnier, J., Estell, D., Miller, J., Graycar, T., Adams, R., & Power, S. (1987) *ACS Symposium Series 334* (LeBaron, H. M., Mumma, R. O., Honeycutt, R. C., & Deusing, J. H., Eds.) pp 139–147, American Chemical Society, Washington, DC.

Brown, K. A., Brick, P., & Blow, D. M. (1987) *Nature 326*, 416–418.

Bryan, P. N., Rollence, M.-L., Pantoliano, M. W., Wood, J., Finzel, B. C., Gilliland, G. L., Howard, A. J., & Poulos, T. L. (1987) *Proteins: Struct., Funct., Genet. 1*, 326–334.

Carter, P., & Wells, J. A. (1987) *Science 237*, 394.

Carter, P., & Wells, J. A. (1988) *Nature 332*, 564–568.

Carter, P., & Wells, J. A. (1990) *Proteins: Struct., Funct., Genet.* (in press).

Carter, P., Nilsson, B., Burnier, J. P., Burdick, D., & Wells, J. A. (1989) *Proteins: Struct., Funct., Genet. 6*, 240–248.

Carter, P. J., Winter, G., Wilkinson, A. J., & Fersht, A. R. (1984) *Cell 38*, 835–840.

Chothia, C., & Lesk, A. (1986) *EMBO J. 5*, 823–826.

Craik, C. S., Largman, C., Fletcher, T., Roczniak, S., Barr, P. J., Fletterick, R., & Rutter, W. J. (1985) *Science 228*, 291–297.

Cunningham, B. C., & Wells, J. A. (1987) *Protein Eng. 1*, 319–325.

Cunningham, B. C., Henner, D. J., & Wells, J. A. (1990) *Science 247*, 1461–1465.

Ebright, R. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 303–307.

Ebright, R. H., Kolb, A., Buc, H., Kunkel, T. A., Krakow, J. S., & Beckwith, J. (1987) *Proc. Natl. Acad. sci. U.S.A. 84*, 6083–6087.

Empie, M. W., & Laskowski, M., Jr. (1982) *Biochemistry 21*, 2274–2284.

Fersht, A. (1985) in *Enzyme Structure and Mechanism*, 2nd ed., Chapters 3, 12, and 13, W. H. Freeman and Co., New York.

Fersht, A. R. (1987) *Biochemistry 26*, 8031–8037.

Fersht, A. R., Wilkinson, A. J., Carter, P., & Winter, G. (1985) *Biochemistry 24*, 5858–5861.

Hecht, M. H., Sturtevant, J. M., & Sauer, R. T. (1986) *Proteins: Struct., Funct., Genet. 1* 43–46.

Horovitz, A., & Rigbi, M. (1985) *J. Theor. Biol. 116*, 149–159.

Howell, E. E., Villafranca, J. E., Warren, M. S., Oatley, S. J., & Kraut, J. (1986) *Science 231*, 1123–1128.

Hurle. M. R., Tweedy, N. B., & Matthews, C. R. (1986) *Biochemistry 25*, 6356–6360.

Imanaka, T., Shibazaki, M., & Takagi, M. (1986) *Nature 324*, 695–697.

Jencks, W. P. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4046–4050.

Jones, M. D., Lowe, D. M., Borgford, T., & Fersht, A. R. (1986) *Biochemistry 25*, 1887–1891.

Katz, B. A., & Kossiakoff, A. (1986) *J. Biol. Chem. 261*, 15480–15485.

Laskowski, M., Jr., Tashiro, M., Empie, M. W., Park, S. J., Kato, I., Ardelt, W., & Mieczorek, M. (1983) in *Protease Inhibitors: Medical and Biological Aspects* (Katunuma, N., Ed.) pp 55–68, Japan Scientific Societies Press, Tokyo, Japan.

Laskowski, M., Jr., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M. W., Kohr, W. J., Park, S. J., Parks, K., Schatzley, B. L., Schoenberger, O. L., Tashiro, M., Vichot, G., Whatley, H. E., Wieczorek, A., & Wieczorek, M. (1987) *Biochemistry 26*, 202–221.

Laskowski, M., Jr., Park, S. J., Tashiro, M., & Wynn, R. (1989) in *Protein Recognition of Immobilized Ligands*, UCLA Symposia on Molecular and Cellular Biology (Hutchens, T. W., Ed.) Vol. 80, pp 149–160, A. R. Liss, New York.

Leatherbarrow, R. J., Fersht, A. R., & Winter, G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7840–7844.

Lehming, N., Sartorius, J., Kisters-Woike, B., von Wilcken-Bergmann, B., & Muller-Hill, B. (1990) *EMBO J. 9*, 615–621.

Liao, H., McKenzie, T., & Hageman, R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 83*, 576–580.

Lowe, D. M., Fersht, A. R., Wilkinson, A. J., Carter, P., & Winter, G. (1985) *Biochemistry 24*, 5106–5109.

Matsumura, M., Yasumura, S., & Aiba, S. (1986) *Nature 323*, 356–358.

Matsumura, M., Signor, G., & Matthews, B. W. (1989) *Nature 342*, 291–294.

Matthews, B. W. (1987) *Biochemistry 26*, 6885–6888.

Mayer, R. J., Chen, J.-T., Taira, K., Fierke, C. A., & Benkovic, S. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7718–7720.

Mitchinson, C., & Baldwin, R. L. (1986) *Proteins: Struct., Funct., Genet. 1*, 23–33.

Nelson, H. C. M., & Sauer, R. T. (1985) *Cell 42*, 549-558.
Otlewski, J., & Laskowski, M., Jr. (1990) (submitted for publication).
Pantoliano, M. W., Whitlow, M., Wood, J. F., Dodd, S. W., Hardman, K. D., Rollence, M. L., & Bryan, P. N. (1989) *Biochemistry 28*, 7205-7213.
Russell, A. J., & Fersht, A. R. (1987) *Nature 328*, 496-500.
Sandberg, W. S., & Terwilliger, T. C. (1989) *Science 245*, 54-57.
Sarai, A., & Takeda, Y. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6513-6517.
Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157-162.
Scrutton, N. S., Berry, A., & Perham, R. N. (1990) *Nature 343*, 38-43.
Shortle, D., & Meeker, A. K. (1986) *Proteins: Struct., Funct., Genet. 1*, 81-89.
Takeda, Y., Sarai, A., & Rivera, V. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 439-443.
Wells, J. A., & Estell, D. A. (1988) *Trends Biochem. Sci. 13*, 291-297.
Wells, J. A., Cunningham, B. C., Graycar, T. P., & Estell, D. A. (1987a) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5167-5171.
Wells, J. A., Powers, D. B., Bott, R. R., Graycar, T. P., & Estell, D. A. (1987b) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1219-1223.
Wells, J. A., Cunningham, B. C., Graycar, T. P., Estell, D. A., & Carter, P. (1987c) *Cold Spring Harbor Symp. Quant. Biol. 52*, 647-652.
Wetzel, R., Perry, L. J., Baase, W. A., & Becktel, W. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 875*, 401-405.
Wilde, J. A., Bolton, P. H., Dell'Acqua, M., Hibler, D. W., Pourmotabbed, T., & Gerlt, J. A. (1988) *Biochemistry 27*, 4127-4132.
Wilkinson, A. J., Fersht, A. R., Blow, D. M., & Winter, G. (1983) *Biochemistry 22*, 3581-3586.
Wilkinson, A. J., Fersht, A. R., Blow, D. M., Carter, P., & Winter, G. (1984) *Nature 307*, 187-188.

Accelerated Publications

Role of Tyrosine M210 in the Initial Charge Separation of Reaction Centers of *Rhodobacter sphaeroides*†

Ulrich Finkele, Christoph Lauterwasser, and Wolfgang Zinth
Physik Department der Technischen Universität, D-8000 München 2, FRG
Kevin A. Gray and Dieter Oesterhelt*
Max-Planck-Institut für Biochemie, D-8033 Martinsried, FRG
Received May 22, 1990; Revised Manuscript Received July 12, 1990

ABSTRACT: Femtosecond spectroscopy was used in combination with site-directed mutagenesis to study the influence of tyrosine M210 (YM210) on the primary electron transfer in the reaction center of *Rhodobacter sphaeroides*. The exchange of YM210 to phenylalanine caused the time constant of primary electron transfer to increase from 3.5 ± 0.4 ps to 16 ± 6 ps while the exchange to leucine increased the time constant even more to 22 ± 8 ps. The results suggest that tyrosine M210 is important for the fast rate of the primary electron transfer.

The primary photochemical event during photosynthesis of bacteriochlorophyll- (Bchl-) containing organisms is a light-induced charge separation within a transmembrane protein complex called the reaction center (RC). The crystal structures of RC's from *Rhodopseudomonas (Rps.) viridis* and *Rhodobacter (Rb.) sphaeroides* have been solved to high resolution [reviewed in Deisenhofer and Michel (1989), Chang et al. (1986), Tiede et al. (1988), and Rees et al. (1989)]. The RC from *Rb. sphaeroides* contains three protein subunits referred to as L, M, and H, according to their respective mobilities in SDS-polyacrylamide gels. Associated with the L and M subunits are the cofactors, consisting of four Bchl *a*, two bacteriopheophytin (Bph) *a*, one atom of non-heme ferrous iron, two quinones (Q_A and Q_B), and in some species one carotenoid [reviewed in Parson (1987) and Feher et al. (1989)]. The cofactors are arranged in two branches (Figure 1) with an approximate C_2 axis of symmetry. The kinetic data support a model in which the primary electron transfer proceeds after light absorption by the primary donor [a special pair of Bchl referred to as P; reviewed in Kirmaier and Holten (1987)]. The absorption of light generates the excited electronic state P^*, which has a lifetime of approximately 3 ps. An electron is transferred from P along only one branch (the so-called A-branch). It is generally accepted that after approximately 3 ps the electron arrives at the Bph on the A-side (H_A) and after 220 ps it reaches Q_A. The role of the accessory Bchl located between P and H_A (referred to as B_A) has not been definitely assigned. Recently, we have shown that at room temperature an additional kinetic (τ = 0.9 ps) component is detectable (Holzapfel et al., 1989). The spectral properties and the kinetic constants lead to the conclusion that the corresponding intermediate is the radical pair $P^+B_A^-$ (Holzapfel et al., 1990)

† Financial support was from the Deutsche Forschungsgemeinschaft,

NOTICE OF OFFICE PLAN TO CEASE SUPPLYING COPIES OF CITED U.S. PATENT REFERENCES WITH OFFICE ACTIONS, AND PILOT TO EVALUATE THE ALTERNATIVE OF PROVIDING ELECTRONIC ACCESS TO SUCH U.S. PATENT REFERENCES

Summary

The United States Patent and Trademark Office (Office or USPTO) plans in the near future to: (1) cease mailing copies of U.S. patents and U.S. patent application publications (US patent references) with Office actions except for citations made during the international stage of an international application under the Patent Cooperation Treaty and those made during reexamination proceedings; and (2) provide electronic access to, with convenient downloading capability of, the US patent references cited in an Office action via the Office's private Patent Application Information Retrieval (PAIR) system which has a new feature called "E-Patent Reference." Before ceasing to provide copies of U.S. patent references with Office actions, the Office shall test the feasibility of the E-Patent Reference feature by conducting a two-month pilot project starting with Office actions mailed after December 1, 2003. The Office shall evaluate the pilot project and publish the results in a notice which will be posted on the Office's web site (www.USPTO.gov) and in the Patent Official Gazette (O.G.). In order to use the new E-Patent Reference feature during the pilot period, or when the Office ceases to send copies of U.S. patent references with Office actions, the applicant must: (1) obtain a digital certificate from the Office; (2) obtain a customer number from the Office, and (3) properly associate applications with the customer number. The pilot project does not involve or affect the current Office practice of supplying paper copies of foreign patent documents and non-patent literature with Office actions. Paper copies of references will continue to be provided by the USPTO for searches and written opinions prepared by the USPTO for international applications during the international stage and for reexamination proceedings.

Description of Pilot Project to Provide Electronic Access to Cited U.S. Patent References

On December 1, 2003, the Office will make available a new feature, E-Patent Reference, in the Office's private PAIR system, to allow more convenient downloading of U.S. patents and U.S. patent application publications. The new feature will allow an authorized user of private PAIR to download some or all of the U.S. patents and U.S. patent application publications cited by an examiner on form PTO-892 in Office actions, as well as U.S. patents and U.S. patent application publications submitted by applicants on form PTO/SB08 (1449) as part of an IDS. The retrieval of some or all of the documents may be performed in one downloading step with the documents encoded as Adobe Portable Document format (.pdf) files, which is an improvement over the current page-by-page retrieval capability from other USPTO systems.

Steps to Use the New E-Patent Reference Feature During the Pilot Project and Thereafter

Access to private PAIR is required to utilize E-Patent Reference. If you don't already have access to private PAIR, the Office urges practitioners, and applicants not represented by a practitioner, to take advantage of the transition period to obtain a no-cost USPTO Public Key Infrastructure (PKI) digital certificate, obtain a USPTO customer number, associate all of their pending and new application filings with their customer number, install no-cost software (supplied by the Office) required to access private PAIR and E-Patent Reference feature, and make appropriate arrangements for Internet access. The full instructions for obtaining a PKI digital certificate are available at the Office's Electronic Business Center (EBC) web page at: <http://www.uspto.gov/ebc/downloads.html>. Note that a notarized signature will be required to obtain a digital certificate.

To get a Customer Number, download and complete the Customer Number Request form, PTO-SB125, at: http://www.uspto.gov/web/forms/sb0125.pdf. The completed form can then be transmitted by facsimile to the Electronic Business Center at (703) 308-2840, or mailed to the address on the form. If you are a registered attorney or patent agent, then your registration number must be associated with your customer number. This is accomplished by adding your registration number to the Customer Number Request form. A description of associating a customer number with an application is described at the EBC web page at: http://www.uspto.gov/ebc/registration_pair.html.

The E-Patent Reference feature will be accessed using a new button on the private PAIR screen. Ordinarily all of the cited U.S. patent and U.S. patent application publication references will be available over the Internet using the Office's new E-Patent Reference feature. The size of the references to be downloaded will be displayed by E-Patent Reference so the download time can be estimated. Applicants and registered practitioners can select to download all of the references or any combination of cited references. Selected references will be downloaded as complete documents as Adobe Portable Document Format (.pdf) files. For a limited period of time, the USPTO will include a copy of this notice with Office actions to encourage applicants to use this new feature and, if needed, to take the steps outlined above in order to be able to utilize this new feature during the pilot and thereafter.

During the two-month pilot, the Office will evaluate the stability and capacity of the E-Patent Reference feature to reliably provide electronic access to cited U.S. patent and U.S. patent application publication references. While copies of U.S. patent and U.S. patent application publication references cited by examiners will continue to be mailed with Office actions during the pilot project, applicants are encouraged to use the private PAIR and the E-Patent Reference feature to electronically access and download cited U.S. patent and U.S. patent application publication references so the Office will be able to objectively evaluate its performance. The public is encouraged to submit comments to the Office on the usability and performance of the E-Patent Reference feature during the pilot. Further, during the pilot period registered practitioners, and applicants not represented by a practitioner, are encouraged to experiment with the feature, develop a proficiency in using the feature, and establish new internal processes for using the new access to the cited U.S. patents and U.S. patent application publications to prepare for the anticipated cessation of the current Office practice of supplying copies of such cited

references. The Office plans to continue to provide access to the E-Patent Reference feature during its evaluation of the pilot.

Comments

Comments concerning the E-Patent Reference feature should be in writing and directed to the Electronic Business Center (EBC) at the USPTO by electronic mail at eReference@uspto.gov or by facsimile to (703) 308-2840. Comments will be posted and made available for public inspection. To ensure that comments are considered in the evaluation of the pilot project, comments should be submitted in writing by January 15, 2004.

Comments with respect to specific applications should be sent to the Technology Centers' customer service centers. Comments concerning digital certificates, customer numbers, and associating customer numbers with applications should be sent to the Electronic Business Center (EBC) at the USPTO by facsimile at (703) 308-2840 or by e-mail at EBC@uspto.gov.

Implementation after Pilot

After the pilot, its evaluation, and publication of a subsequent notice as indicated above, the Office expects to implement its plan to cease mailing paper copies of U.S. patent references cited during examination of non provisional applications on or after February 2, 2004; although copies of cited foreign patent documents, as well as non-patent literature, will still be mailed to the applicant until such time as substantially all applications have been scanned into IFW.

For Further Information Contact

Technical information on the operation of the IFW system can be found on the USPTO website at http://www.uspto.gov/web/patents/ifw/index.html. Comments concerning the E-Patent Reference feature and questions concerning the operation of the PAIR system should be directed to the EBC at the USPTO at (866) 217-9197. The EBC may also be contacted by facsimile at (703) 308-2840 or by e-mail at EBC@uspto.gov.

Date. 12/1/03

Nicholas P. Godici

Nicholas P. Godici
Commissioner for Patents

USPTO TO PROVIDE ELECTRONIC ACCESS TO CITED U.S. PATENT REFERENCES WITH OFFICE ACTIONS AND CEASE SUPPLYING PAPER COPIES

In support of its 21st Century Strategic Plan goal of increased patent e-Government, beginning in June 2004, the United States Patent and Trademark Office (Office or USPTO) will begin the phase-in of its E-Patent Reference program and hence will: (1) **provide downloading capability of the U.S. patents and U.S. patent application publications cited in Office actions** via the E-Patent Reference feature of the Office's Patent Application Information Retrieval (PAIR) system; and (2) **cease mailing paper copies of U.S. patents and U.S. patent application publications with Office actions** (in applications and during reexamination proceedings) except for citations made during the international stage of an international application under the Patent Cooperation Treaty (PCT). In order to use the new E-Patent Reference feature applicants must: (1) obtain a digital certificate and software from the Office; (2) obtain a customer number from the Office; and (3) properly associate patent applications with the customer number. Alternatively, copies of all U.S. patents and patent application publications can be accessed without a digital certificate from the USPTO web site, from the USPTO Office of Public Records, and from commercial sources. The Office will continue the practice of supplying paper copies of foreign patent documents and non-patent literature with Office actions. Paper copies of cited references will continue to be provided by the USPTO for international applications during the international stage.

Schedule

June 2004	TCs 1600, 1700, 2800 and 2900
July 2004	TCs 3600 and 3700
August 2004	TCs 2100 and 2600

All U.S. patents and U.S. patent application publications are available on the USPTO web site. However, a simple system for downloading the cited U.S. patents and patent application publications has been established for applicants, called the E-Patent Reference system. As E-Patent Reference and Private PAIR require participating applicants to have a customer number, retrieval software and a digital certificate, all applicants are strongly encouraged to contact the Patent Electronic Business Center to acquire these items. To be ready to use this system by June 1, 2004, contact the Patent EBC as soon as possible by phone at 866-217-9197 (toll-free), 703-305-3028 or 703-308-6845 or electronically via the Internet at ebc@uspto.gov.

Other Options

The E-Patent Reference function requires the applicant to use the secure Private PAIR system, which establishes confidential communications with the applicant. Applicants using this facility must receive a digital certificate, as described above. Other options for obtaining patents which do not require the digital certificate include the USPTO's free Patents on the Web program (http://www.uspto.gov/patft/index.html). The USPTO's Office of Public Records also supplies copies of patents for a fee (http://ebiz1.uspto.gov/oems25p/index.html). Commercial sources also provide U.S. patents and patent application publications.

For complete instructions see the Official Gazette Notice, USPTO TO PROVIDE ELECTRONIC ACCESS TO CITED U.S. PATENT REFERENCES WITH OFFICE ACTIONS AND CEASE SUPPLYING PAPER COPIES, on the USPTO web site.